DUKE
UNIVERSITY

LIBRARY

UKE UNIVERSITY RESEARCH STUDIES IN EDUCATION . NUMBER 5

THE EFFECT OF VARIED AMOUNTS OF PHONETIC TRAINING ON PRIMARY READING

DUKE UNIVERSITY LIBRARY
DURHAM N. C.

BY

DONALD C. AGNEW

DUKE UNIVERSITY PRESS
Durham, N. C.
1939

378.756
D877R
no.5

DUKE UNIVERSITY RESEARCH STUDIES IN EDUCATION

THE EFFECT OF VARIED AMOUNTS OF PHONETIC TRAINING ON PRIMARY READING

THE EFFECT OF VARIED AMOUNTS OF PHONETIC TRAINING ON PRIMARY READING

BY

DONALD C. AGNEW
Coker College, Hartsville, South Carolina

DUKE UNIVERSITY PRESS
Durham, N. C.
1939

COPYRIGHT, 1939, BY THE
DUKE UNIVERSITY PRESS

PRINTED IN THE UNITED STATES OF AMERICA BY
THE SEEMAN PRINTERY, INC., DURHAM, N. C.

FOREWORD

For one reason or another, research has not yielded an unequivocal answer to the place of phonetic training in the teaching of primary reading. Sometimes experiments have covered but a short period of time; sometimes programs of measurement by means of which the effects of instruction have been assessed have been unfortunately limited and have neglected important outcomes in reading; and sometimes evaluations have been based upon immediate results only, with little consideration of deferred results of possible value. Meanwhile, in the absence of consistent and competent research findings the educational theorist and the practical teacher have perforce continued to follow their own experience, their insights, and their prejudices.

Dr. Agnew has wisely not attempted a final answer to the questions he has investigated. He has not sought, on the basis of his research data, to say whether phonetic training should be emphasized or minimized in primary reading, for he has recognized that a decision in this matter rests only in part upon data such as he reports, and rests much more upon one's conception of the purposes and aims of reading as a whole. What Dr. Agnew has done is to collect important new data on the effects of large and of small amounts of phonetic training. In so doing, he has employed a measurement program far more comprehensive than has usually characterized studies of phonetic instruction, and he has included practically all aspects of reading ability as this ability has been analyzed by others.

The research herein reported should be of interest not alone to the teacher and the administrator for its practical results and its implications; it should be of interest as well to the student of educational research. Dr. Agnew has used a number of procedures, involving both statistical method and measuring techniques, which may be of considerable worth in the investigation of other instructional problems of the school.

William A. Brownell.

ACKNOWLEDGMENTS

I am deeply indebted to Dr. William A. Brownell, Professor of Educational Psychology, Duke University, for his proposal of the thesis topic, for his aid and guidance in carrying on the investigations, for his suggestions as to the treatment of the data, and for his critical advice with respect to reporting the findings. I am indebted also to Dr. Howard Easley, Assistant Professor of Education, Duke University, for his suggestions concerning the statistical treatment of the results and concerning certain psychological interpretations. I am grateful to Dr. Holland Holton, Head of the Department of Education, Duke University, for his encouragement and aid throughout the investigations.

The investigations were made possible through the generous cooperation of Superintendent P. S. Daniel and Assistant Superintendent, Miss Mildred English, of the Raleigh, N. C., Schools; and of Superintendent W. F. Warren and Mrs. J. A. Robinson (Primary Supervisor), of the Durham, N. C., Schools. The elementary school principals and the teachers in Grades I, II, and III, in both Raleigh and Durham, were very helpful throughout the testing programs. I appreciate, too, the valuable assistance of the members of the classes in Experimental Education and in Educational Measurements (Duke University), who gave numerous tests in the Raleigh and Durham schools.

DONALD C. AGNEW.

Coker College
October 11, 1938

TABLE OF CONTENTS

CHAPTER		PAGE
I.	INTRODUCTION	3
II.	A BRIEF ACCOUNT OF THE INVESTIGATIONS TO BE REPORTED	9
	A. The particular problems investigated	9
	B. The general theory of the investigations	10
	C. A brief account of the procedure employed	12
	D. The measures used in the investigations	12
	E. The Durham investigation	14
III.	THE TECHNIQUE OF THE FIRST INVESTIGATION	15
	A. Selection of subjects	15
	B. The measures of phonetic experience	15
	C. The testing program	18
IV.	THE METHODS OF TREATING THE RESULTS	21
	A. Comparison based on the Gross Phonetic Experience Scores	21
	B. Comparisons of patterns of phonetic experience	22
V.	RESULTS AND CONCLUSIONS OF THE RALEIGH INVESTIGATION	27
VI.	THE PURPOSE AND TECHNIQUE OF THE DURHAM INVESTIGATION	35
	A. Purpose	35
	B. The technique of the investigation	35
VII.	RESULTS AND CONCLUSIONS OF THE DURHAM INVESTIGATION	40
	Conclusions	42
VIII.	BRIEF SUMMARY OF THE RESULTS OF THE INVESTIGATIONS, AND CERTAIN EDUCATIONAL IMPLICATIONS	44
	A. The relation of the investigations to the controversial issues with regard to phonetic instruction	44
	B. Theoretical considerations	45
	C. Educational implications	47
BIBLIOGRAPHY		48

TABLES

TABLE		PAGE
1.	Studies of the Phonetic Content of English Vocabularies at the Primary Level	4
2.	Studies of the Value of Phonetic Training in the Elementary Grades	5
3.	Distributions of Scores on the Teachers' Blanks by Grades	16
4.	Differences Between the Responses of Teacher F. L. and Teacher R. H. on Ten Items of the Teachers' Blank	17
5.	Distribution of Gross Phonetic Experience Scores (III-A Pupils in Raleigh)	18
6.	Distributions of M.A.'s and I.Q.'s for Groups G_H and G_L	22
7.	Frequency Distribution of Phonetic Experience Scores of Pupils in Grade I	24
8.	Distributions of Phonetic Experience Scores for Grades II and III Combined Before and After Equating	24
9.	Distributions of Equated Groups	25
10.	Summary of the Differences in Terms of the P. E. of the Differences	27
11.	Means and Standard Deviations of the M.A.'s for the Various Patterns	28
12.	Correlations Between Test Scores of Groups G_H and G_L (Groups Matched for M.A. and I.Q.)	29
13.	Frequency Distribution of Gross Phonetic Experience Scores (110 Cases, Durham)	37
14.	Frequency Distributions of M.A.'s and I.Q.'s of the Raleigh and Durham Groups Selected by Pairing	38
15.	Factors in the Composition of the Durham and Raleigh Groups	39
16.	Differences Between the Raleigh and Durham Groups in Terms of the Differences Between the Means of the Test Scores (89 Cases in Each Group)	41

THE EFFECT OF VARIED AMOUNTS OF PHONETIC TRAINING ON PRIMARY READING

CHAPTER I

INTRODUCTION

Statement of problem.—The investigation to be reported was undertaken in an effort to determine the effects of varied amounts of phonetic training on certain reading abilities as measured by a battery of tests. The particular reading abilities measured were those which, it was thought, should shed light on some of the controversial questions that arise in connection with the use of phonetic methods in beginning reading instruction.

The history of phonetic methods of instruction.—The history of phonetic methods of instruction is too long to be recounted here. It should be pointed out, however, that since the beginning of the nineteenth century when Noah Webster's blueback speller stressed the sounding of letters and syllables in reading and spelling, phonetic instruction has had numerous periods of popularity and unpopularity. Furthermore, certain kinds of phonetic instruction, for example, the syllable method, the alphabet method, the word method, and special modifications of these methods, have been used slightly or extensively in certain localities. During the past decade extreme differences of opinion have been expressed by various educators as to the value of phonetic instruction. Thus, the history of phonetic instruction is one of inconsistency and controversy.[1]

Quantitative studies related to phonetic training.—The quantitative studies directly related to the problem may be considered in two categories: studies of the phonetic content of vocabularies, and studies of the value of phonetic training.

The studies of the phonetic content of vocabularies are listed in Table 1. The results of the studies of sound frequency have had considerable bearing on the teaching of phonetics. They have shown, in general, the following results: (a) The English language has a large number of nonphonetic words. (b) Letters and letter-combinations often have a large number of possible sounds. Horn

[1] A more detailed history of phonetic methods of instruction is given in Chapter I in the unpublished thesis by the writer: "The Effect of Varied Amounts of Phonetic Training on Primary Reading" (Duke University, 1936). In the same place, in Chapter II, and in considerable detail in Appendix A, are given an account and a criticism of the quantitative studies related directly and indirectly to the problem. These studies are listed in the Bibliography.

TABLE 1

STUDIES OF THE PHONETIC CONTENT OF ENGLISH VOCABULARIES AT THE PRIMARY LEVEL

Author	Problem	Technique
Atkins (3)*	Constancy of phonetic sounds in the first 2,500 words of the Thorndike list.	Analysis of words and calculation of the frequency of phonetic and nonphonetic sounds in the Thorndike list.
Cordts (11)	Classification of sounds in primary vocabulary.	Analysis of words and calculation of frequency of phonetic sounds in a large number of readers.
Horn (26)	Classification of sounds of the letter "a" in vocabulary of the first three grades.	Analysis of words and calculation of frequency of all sounds of the letter "a" in the vocabulary prepared by Cordts.
Vogel, Jaycox, and Washburn (44)	Classification of list of phonetic elements for Grades I and II.	Analysis of words and calculation of frequency of phonograms in Packer's vocabulary and seventeen primers.
Washburn and Vogel (45)	Classification of phonetic elements in Gates's Primary Reading Vocabulary for Grade II.	Analysis of vocabulary and calculation of frequency of sounds in the Gates list.

*Numbers in parentheses following the names indicate the numbers of the references given in the Bibliography.

(26),[2] for example, finds that the letter "a" has almost fifty different sounds when used alone or in digraphs. He shows also that "ea," for example, may have as many as eight different sound interpretations when found in different settings. (c) Lists of the most frequently occurring sounds have been made so that emphasis may be placed upon the sounds according to their deemed importance. (d) Elementary vocabularies have been analyzed for their phonetic elements.

Various inferences have been drawn from these studies. Some educators, for example, Horn (26), claim that the studies show the teaching of phonetics to be an almost impossible task for the elementary school. Others, for example, Wheat (46), insist that phonetics is necessary for the acquisition of reading skills.

The results of the experimental investigations of the value of phonetic training are by no means conclusive. These studies are listed in Table 2. The organization of the experiments was not such that the studies could yield unequivocal evidence. The available data seem to point tentatively to the following conclusions: (a) Investigators tend to agree that the first few months of reading should emphasize the "look and say" method.[3] Sexton and Her-

[2] Hereafter in this report, numbers in parentheses following the name indicate the number of the reference listed in the Bibliography at the end of the monograph.

[3] Winch (47) holds the opposite opinion.

TABLE 2

STUDIES OF THE VALUE OF PHONETIC TRAINING IN THE ELEMENTARY GRADES

Author	Problem	Technique
Currier (12)*	Value of phonetic training in Grades I and II. Phonetic training given one group for two years.	Control group experiment.
Garrison and Heard (15)	Value of fifteen minutes daily drill in phonetics in Grades I and II. Phonetic training given for two years.	Control group experiment.
Gates (22)	Value of fifteen minutes daily drill in phonetics during about six months of Grade I (two experiments).	Control group experiment.
Mosher (31)	Value of "look and say" method for beginners without phonetic training during the first year.	Comparison of groups of of differing ability.
Mosher and Newhall (32)	Value of "look and say" method *vs.* phonetic methods, Grade I. Phonetic training given for seven months.	Control group experiment.
Sexton and Herron (39)	Value of fifteen minutes daily drill in phonetics in Grade I. Phonetic training given for seven months.	Control group experiment.
Winch (47)	Value of fifteen minutes daily drill in phonetics *vs.* "look and say" methods. Phonetic training given on twenty-five consecutive school days.	Control group experiment.

*See note under Table 1.

ron (39), Gates (22), Mosher and Newhall (32), and Garrison and Heard (15) agree in this conclusion. (b) Gates (22) and Sexton and Herron (39) found the "intrinsic method" to be somewhat superior to the special drill periods, at least when the results are measured after a brief period of time. (c) It seems to be conceded that phonetic training aids in word recognition. (d) Garrison and Heard (15) found that such training, however, inhibits fluency in oral reading. (e) Currier (12) found that some children need phonetic training more than do others.

A review of the studies indicates that the evidence on the value of phonetic training is limited. The need for further investigation is apparent.[4]

Difference of opinion among teachers.—Nearly every elementary-school teacher has deep-seated beliefs about the importance or weakness of phonetic methods. Many teachers believe intensive phonetic training desirable. Nila B. Smith points out, on the other hand: "There is an impression among some teachers that phonetics is a disgrace, that this phase of instruction is of no value and is generally being abandoned."[5]

[4] Other studies indirectly related to phonetic training are reviewed in the original thesis. The bibliographical references to these studies are numbered as follows: 4, 6, 9, 13, 21, 24, 28, 30, 31, 41, 42.

[5] From *American Reading Instruction,* 1934, by permission of the author, Nila B. Smith, and the publisher, Silver Burdett Company. See Smith (40), p. 220.

The attitude of textbook writers.—The attitude of textbook writers is more conservative. Smith summarizes the situation as follows:

> While some . . . textbook writers seem less certain than writers of former times on this subject, there is nothing to indicate that any one of them takes the extreme attitude of dispensing with phonetics entirely. Every manual that has appeared in connection with a basal series of readers during this period (1925-1934) has recognized phonetics. Various states of confidence in the value of phonetics are expressed by authors, but they all discuss this phase of reading and outline procedures for teaching it.[6]

Disagreement among educators.—McKee summarizes the situation with respect to the attitude of educators toward the problem as follows:

> The question of instruction in phonics has aroused a great deal of controversy. Some educators have held to the proposition that phonetic training is not only futile and wasteful but also harmful to the best interests of a reading program. Others believe that since the child must have some means of attacking strange words instruction in phonics is imperative. There have been disputes also relative to the amount of phonics to be taught, the time when the teaching should take place, and the methods to be used. In fact the writer knows of no problem around which more disputes have centered.[7]

Examples of these differences of opinion are numerous in the literature. O'Brien[8] is highly in favor of phonetic training. Wheat holds the same opinion. He says: "Phonic analysis is the device necessary to train pupils to avoid periods of confusion."[9]

On the other hand, Gates considers phonetic training of doubtful value, particularly in the first grade, and especially when taught by the traditional methods. He says:

> The great mistake in American teaching has been the assumption that phonetic skill was all-important and sufficient, that the other types of training could be neglected, and that the more phonetics the pupil got the better. These mistakes have resulted not only in waste but frequently in the production of a special type of difficulty in reading. So excessive has phonetic drill often been that pupils have become not only "word-form conscious" at the expense of interest in meanings, but, even worse, they have become "word detail conscious" . . . . Thus phonetic skill in moderation is useful; in less degree, it leaves the pupil handicapped; in greater degree it may result in a more serious deficiency.[10]

[6] *Ibid.*
[7] McKee (29), p. 191. Quoted by special permission of the author.
[8] O'Brien (33), p. 225.
[9] Wheat (46), p. 199.
[10] Gates (16), pp. 125-126. Quoted by special permission of the publishers: Macmillan Co.

Cabell[11] agrees with Gates and contends that the teaching of phonetics is largely on the defensive.

The quotations cited are only a few of the many diverse opinions that could be found, but they are enough to bear out the statement taken from McKee. In the following paragraphs, a summary of the arguments that are advanced for and against phonetic training is given.

The case in favor of phonetics.—The following arguments may be said to sum up the case in favor of phonetic training:

1. Phonetic training has had a long history; during this period of years until quite recently, it has been provided in increasingly large amounts. Procedures that have been used in the teaching of reading for a century should be scrutinized very carefully before they are abandoned.
2. Phonetic training gives the pupil independence in recognizing words previously learned. This ability becomes steadily more important in connection with silent reading.
3. Phonetic training aids in "unlocking" new words by giving the pupil a method of sound analysis.
4. Phonetic training encourages correct pronunciation and enunciation.
5. Phonetic training gives valuable "ear training" in recognizing and differentiating sounds.
6. Phonetic training improves the quality of oral reading, for instance, in breath control and in speech co-ordination.
7. Phonetic training improves spelling.
8. Phonetic training is a valuable background for shorthand.
9. Many cases of reading disability may be traced to deficiencies in word recognition and sound analysis. These disabilities are often overcome by remedial procedures involving phonetic training.

The case against phonetics.—The disadvantages attributed to phonetic training may be summarized as follows:

1. Phonetic training tends to isolate words from their meaningful function by emphasizing sound.
2. Phonetic training tends to lead to the neglect of context clues.
3. Phonetic training tends to sacrifice interest in the content of reading.

[11] Cabell (7), pp. 370-373.

4. Phonetic training leads to unnecessarily laborious recognition of familiar words.
5. Phonetic training is impractical because of the nonphonetic character of English.
6. Phonetic training is unnecessary for many pupils since its advantages can be obtained without formal training.
7. Phonetic training encourages the breaking of words into unnecessarily small units.
8. Phonetic training narrows the eye-voice span.
9. Phonetic training tends to emphasize too explicit articulation.

Need for experimental evidence.—The claims and objections listed above are inferences based largely on a priori considerations. As such they are at best tentative rather than final. Does phonetic training result in these outcomes? The answer must come from scientific data rather than from mere speculation. Perhaps phonetic training is neither as bad as one group claims nor as good as the other group insists. The studies reported here were undertaken in an effort to help solve some of the problems raised in the foregoing paragraphs.

CHAPTER II

A BRIEF ACCOUNT OF THE INVESTIGATIONS TO BE REPORTED

A. THE PARTICULAR PROBLEMS INVESTIGATED

The object of the investigations herein reported was to obtain data concerning the validity of some of the claims and objections to phonetics as pointed out in Chapter I. In fact, the two separate studies were direct outgrowths of an analysis of these claims and objections. The general question raised by the analysis was that of the relative value of phonetic training and of nonphonetic training as a basis for teaching reading abilities. This general question involved certain particular questions.

(a) What is the comparative effect of phonetic and nonphonetic reading instructions on speed and comprehension in silent reading? The advocates of phonetic training claim that instruction which stresses phonetic training aids the pupil by giving him methods of attack on unfamiliar words, thus increasing both speed and comprehension. The opponents of phonetic training claim that this type of word analysis tends to make the pupil "word conscious" or "syllable conscious" and thus slows up reading and renders comprehension more difficult because of overemphasis on small sound units.

(b) What are the effects of phonetic and nonphonetic training on speed and accuracy in oral reading? The argument for phonetic training is that such training leads to recognition of words and thus increases speed, and that it also leads to more accurate pronunciation. On the other hand, the argument against phonetic instruction is that such training slows up oral reading because of emphasis on small sound units and that it leads to less accurate pronunciation because of the nonphonetic character of many English words.

(c) What is the effect of phonetic training on eye-voice span? The opponents of phonetic training claim that to make the pupil "word conscious" is to limit his eye-voice span to single words and parts of words. Too much concern with small units of recognition keeps the eye and the voice together and thus prevents the eye from moving considerably ahead of the voice as in fluent oral reading. Exponents of phonetic training naturally minimize this danger.

(d) What are the effects of phonetic and nonphonetic training on reading vocabulary? This question may be stated thus: Does sound analysis, by giving training in independent word recognition, increase vocabulary, or is vocabulary increased with greater ease by "word whole" and "context" methods?

(e) Finally, does or does not phonetic instruction actually result in greater abilities to use phonetic methods? Opponents of phonetic training believe that the skills required for systematic sound analysis are too complex to be learned with any degree of mastery in the primary grades, but that, on the other hand, simple phonetic methods may be developed by the pupil as he feels a need for such methods, without his having been subjected to formal training.

B. THE GENERAL THEORY OF THE INVESTIGATIONS

The decision having been made to investigate the problems named above, the question of the most suitable technique arose.

The control-group technique, as used in most of the previous experiments, has been open to a number of criticisms: (a) It tends to set up artificial situations in which the instructional materials and methods are often new to the teachers. (b) It allows free play for prejudices on the part of teachers so that they may wittingly or unwittingly motivate learning by one method and impede learning by another method. (c) Often it develops a spirit of competition between the experimental and the check group that makes learning under these conditions different from that of the typical school situation.

It was decided, therefore, in the investigations to be undertaken to avoid these limitations through the use of another technique. The plan was to study the results of different teaching procedures which had been employed in ordinary school situations. That is to say, the plan was to locate one group of children who had been given large amounts of phonetic training and another that had been given small amounts of phonetic training, and then to compare these groups with respect to reading skills.

A number of questions then arose: (a) In what grade should the investigation be made? (b) Where could the investigation be instituted in order to provide a wide variation in phonetic experience? (c) How could the amounts of phonetic experience be measured? In order to orient the reader to the discussion that is to follow, a brief answer to these questions is given in the succeeding paragraphs.

(a) *Grade.*—A major criticism of the previous studies has been that the effects of phonetic training have usually been measured immediately after such training has been given. Little time was allowed to elapse between training and testing, although it may well be that the effects of phonetic training may not be appreciable until somewhat later. Even the end of the second grade (if training is given in the first grade) may be too soon to measure these effects. It seemed advisable, therefore, to make the present investigations in the last half of the third grade. By that time the values of phonetic training (if any) should be apparent and measurable. If no difference in reading ability appeared at that time, it would be necessary to conclude that the claims and objections of both advocates and opponents of phonetic training have been exaggerated. It seemed unlikely that effects of different methods in teaching primary reading would for the first time emerge at a later period in the child's progress.

(b) *Location.*—It was thought that the schools selected for investigation should be large enough to give a good sample of third-grade pupils. On the other hand, the sample should be small enough to permit testing throughout the system so that the whole population could be included, the problem of selection of schools or of sampling the pupil population thus being avoided. It was necessary also to choose a school system having a large number of first-, second-, and third-grade teachers in order to obtain a wide variation in the amounts of phonetic teaching. As will be shown later, the city of Raleigh, North Carolina, provided an excellent location for the first investigation.[1]

(c) *Measures of phonetic training.*—The measurement of the phonetic experience to which pupils had been subjected presented a difficult problem. The children could give no information on this point, and interviews with the teachers could yield no real measure. One could, of course, ask a teacher whether or not she used phonetic methods, but her reply, "not much" or "a lot," would be altogether dependent upon her conception of what these terms mean. In other words, the interview method could not yield a quantitative measure of phonetic training. The scale finally employed is described in detail later.

[1] The city of Durham, North Carolina, provided a good location for the second investigation because it is in many ways comparable to Raleigh.

C. A BRIEF ACCOUNT OF THE PROCEDURE EMPLOYED

The procedures used in the investigations fall logically into three parts: (a) those used to secure data on the pupils' phonetic experience; (b) those employed in testing; and (c) those involved in treating the results. These phases of the investigations will be taken up in order in the succeeding paragraphs.

(a) *Securing data on the pupils' phonetic experience.*—Data were obtained in regard to the phonetic experience of the various pupil subjects by means of two instruments called in this report the Pupils' Blank and the Teachers' Blank. These measures are described in detail later. At this point, it is only necessary to say that the Pupils' Blank was designed to secure a record of the pupils' educational histories in terms of the schools they had attended and of the teachers they had had in the first three years. The teachers thus designated were then asked to fill out the Teachers' Blank which was designed to furnish a quantitative measure of the amount of phonetic instruction they had given to the pupils. The relating of these two sets of data made it possible to secure a measure of the amounts of phonetic experience which each child had had in each grade and in the three grades combined.

(b) *The testing program.*—Tests were carefully selected to measure various reading abilities. Group tests of silent reading abilities and vocabulary, and individual tests of oral reading, word pronunciation, eye-voice span and phonetic abilities were administered. In addition to these tests of reading abilities, a group intelligence test was given. Measures of intelligence were intended to relate the factor of general intelligence to the scores on the reading tests and thus aid in the interpretation of the results. A list of these tests is given in Chapter III, together with detailed descriptions of their content and purpose.

(c) *The treatment of the results.*—Groups of pupils who had been subjected to different amounts of phonetic training were compared in terms of the scores they made on the various tests. In order to determine the effects of phonetic and nonphonetic training during the various times in the grade experience of the first two and one-half years of school, a complex statistical analysis of the data was necessary. This analysis is explained in Chapter IV.

D. THE MEASURES USED IN THE INVESTIGATIONS

In the preceding account of the investigation, mention is made of three types of measures, namely, those secured from (a) the

Pupils' Blank, (b) the Teachers' Blank, and (c) the tests of intelligence and reading ability. In this section, the first two will be described in detail and a list of the tests will be presented.

(a) *The Pupils' Blank.*—The Pupils' Blank was designed to record pupils' phonetic experience. A sample of the blank as filled in for John Smith is given below.

PUPILS' BLANK

Name of the pupil John Smith

Grade	School	Teacher's Name	Phonetic Experience Scores
III A	Lewis	Miss A	45
II B	Wiley	Miss B	50
II A	Murphey	Miss C	40
I B	Murphey	Miss D	39
I A	Murphey	Miss D	39
Gross Phonetic Experience Score			213

Copies of the blank were distributed to all the third-grade pupils n the schools for white children in Raleigh. The children were ısked to fill in the columns under "School" and "Teacher's Name," hus giving the name of the school attended during each half grade ogether with the name of the teacher who had taught them in that ıalf grade.[2] It will be noted that John Smith, in the case given above, vas in Grade III A in the Lewis School and that his teacher in that ;rade was Miss A. In Grade II B, he was in the Wiley School and iis teacher was Miss B, etc.

The last column, "Phonetic Experience Scores," was filled in ater with data obtained from the Teachers' Blank.

(b) *The Teachers' Blanks.*[3]—All the teachers who had taught in ;rades I, II, and III during the years 1929-32, and who could be eached in January, 1932, were given one or more Teachers' Blanks. 'hese blanks were designed to secure information concerning the mount of phonetic instruction a teacher had given in each year overed by the blank. One teacher, for example, had taught in ;rade I one year and in Grade II another year. This teacher filled ut a blank for each year. Thus each teacher filled out a maximum f three blanks.

[2] In case a child was unable to furnish the information, it was obtained om the school office.

[3] For a copy of the Teachers' Blank, see Agnew (2), Appendix B. This lank was devised by Dr. William A. Brownell, Professor of Educational sychology, Duke University.

The Teachers' Blank is composed of twenty-five questions, each of which has four possible answers, so worded that they indicate varying degrees of emphasis on phonetic instruction. For instance, Item 6 reads as follows: "With respect to consonant blends (tr, bl, st, etc.) I teach children (a) a very great many, (b) all the common ones, (c) only a few of the most common, (d) none at all as such." The teacher answers the question by placing the letter "a," "b," "c," or "d" in the space for this purpose. In this case, the answer "a" indicates thorough instruction in the phonetic method with respect to this one item, namely, consonant blends, while an answer "d" indicates practically no such instruction, and "b" and "c" represent intermediate degrees. For purposes of scoring, "a" was here given a weight of 4, "b" a weight of 3, "c" a weight of 2, and "d" a weight of 1. All of the other twenty-four questions were similarly constructed and similarly scored. These twenty-five item-scores therefore constituted a scale with a possible range of 25 to 100, 25 representing the least possible amount of phonetic instruction.[4]

The scores obtained from the Teachers' Blanks were transferred to the Pupils' Blank to yield a quantitative measure of the pupils' phonetic experience. Thus, in the case of John Smith,[5] the score of Miss A on the Teachers' Blank was 45, of Miss B, 50, etc. The sum of the scores is called, in this report, the Gross Phonetic Experience Score. This score for John Smith was 213.

(c) *The measures of abilities.*—The abilities which were measured have been pointed out. Descriptions of the tests used to measure these abilities are given in Chapter III.

E. THE DURHAM INVESTIGATION

A wide variation in amounts of phonetic training was found in Raleigh. Nevertheless, in order to test the results of the first experiment, and to secure more data, principally for children who had had a consistently larger amount of phonetic training than the Raleigh children had had, it was felt that a check investigation was desirable. Consequently, a second study was made in Durham where the policy for the three years (1932-1935) had been one of relatively more emphasis on phonetics than was the case in Raleigh. Since the technique in the Durham experiment was essentially the same as that used in Raleigh, the details are not given here. A full account of this investigation is given in Chapter VII.

[4] The reliability of the scores on the Teachers' Blank (obtained by the split-half method, odds *vs.* evens) was found, for 60 cases, to be .96. This remarkably high coefficient is indicative of the inner consistency of the items.

[5] See Pupils' Blank above.

CHAPTER III

THE TECHNIQUE OF THE FIRST INVESTIGATION

A. SELECTION OF SUBJECTS

Number.—For the first selection of pupils, Pupils' Blanks (the nature of which is discussed in the preceding chapter) were distributed to the pupils in Grade III A (the lower half of the third grade). Blanks were secured from 356 pupils. Since, in this investigation, only those pupils were used who had had all their previous schooling in the city of Raleigh, certain eliminations were necessary to preserve the homogeneity of the subjects. (a) All pupils who had repeated grades due to failure or for any other cause were eliminated. They numbered 23. (b) All pupils who, for a term or more, had attended school in any system other than Raleigh were omitted. These pupils numbered 50. (c) Pupils who had been accelerated were omitted. These numbered 13. (d) Ten were omitted because of some ambiguity in their records. (e) Nearly all pupils taught in Grades I to III by teachers who in the period 1929-31 had left the Raleigh system were also omitted.[1] These pupils numbered 30. Thus, in all, 126 pupils were eliminated. After these eliminations were completed, there remained 230 III-A pupils.

Homogeneity of subjects.—These 230 pupils who were finally selected were similar in several respects: (a) They had had all of their school training in the Raleigh schools. (b) This fact assured approximately the same course of study and instructional material. (c) The pupils had all made normal progress in school; that is, they had been neither retarded nor accelerated.

It should be pointed out that this sample of 230 pupils, therefore, represents the whole school population of Grade III A. This fact is of importance to this study, for, while in some of the comparisons to be reported, the number of cases involved is necessarily small, these cases really represent a much larger population.

B. THE MEASURES OF PHONETIC EXPERIENCE

The distribution of scores on the Teachers' Blanks.—The distribution of the scores on the Teachers' Blanks is given in Table 3. Since it was supposed to be the policy in the Raleigh schools not

[1] In the case of one such teacher, it was possible to obtain the necessary data. Her pupils could therefore be retained.

TABLE 3

DISTRIBUTION OF SCORES ON THE TEACHERS' BLANKS BY GRADES*

Scores on the Teachers' Blanks	FREQUENCIES BY GRADES			
	Grade I	Grade II	Grade III	All Grades
75-79	0	1	1	2
70-74	0	3	1	4
65-69	2	3	0	5
60-64	7	3	3	13
55-59	3	0	3	6
50-54	0	2	2	4
45-49	1	2	1	4
40-44	8	6	5	19
35-39	0	1	2	3
30-34	3	3	2	8
Total	24	24	20	68**

*Table 3 should be interpreted as follows: No teacher in the first grade taught phonetics to the extent represented by a score between 75 and 79. One teacher in Grade II and one teacher in Grade III taught phonetics to this extent, etc.

**The total number of teachers noted here exceeds the actual number of teachers because some teachers filled out more than one blank. There were actually 51 teachers. The teachers who filled out more than one blank changed grades during the three years. This fact has no effect on the data because separate measures were obtained for each grade taught.

to teach by phonetic methods, it is interesting to note the wide variation in the scores. The scores range from between 30 and 35 to between 75 and 80. Thus, the scores represent a range of 45 points out of a possible 75. It should be borne in mind that the extreme upper range of possible scores, from 79 to 100, is not represented in the distribution of scores.[2]

A consideration of the frequencies for all grades reveals the distinct bimodal character of the distribution. The difference between the modes is 20, which represents a significant variation in the total range of 45.

The meaning of the scores on the Teachers' Blanks.—It is important to note the facts represented in the scores on the Teachers' Blanks. A teacher's reaction to a single item may be no index to that teacher's general practice with respect to phonetics, but her reactions to the sum of 25 such items very probably does represent her instructional procedure. Differences in teachers' scores may, therefore, be taken to indicate differences in their practices. In order to illustrate the differences in meaning between two scores, the responses of two teachers on ten items of the test are given in Table 4. Teacher

[2] The fact that this extreme was not sampled was one of the reasons for making the second investigation.

TABLE 4

DIFFERENCES BETWEEN THE RESPONSES OF TEACHER F. L. AND TEACHER R. H. ON TEN ITEMS OF THE TEACHERS' BLANK

The nature of the training involved in the item	The Teacher's Response	
	F. L. (Opponent of phonetics. Score: 35)	R. H. (Advocate of phonetics. Score: 75)
1. Phonetic training in relation to sight vocabulary	None at all	Before any words have been learned as sight words.
2. Ear training	Practically never	Regularly.
3. Separate consonant sounds	None at all	All of them.
4. Separate vowel sounds	None at all	All the sounds of all the vowels.
5. Suffixes	None at all	All of them possible.
6. Prefixes	None at all	All of them possible.
7. "Families" of sounds	Pay no attention to them	Identify all words possible by this device.
8. The sounding of individual letters and combinations in new words found in reading	None at all	Regularly before the word is pronounced the first time.
9. Rules for pronunciation	Pay no attention to them	Teach a complete list and require memorization.
10. Teach the sounds of letters by telling stories	Never	Regularly.

F. L. received a score of 35, while teacher R. H. received a score of 75. It is apparent that F. L. gave little instruction in phonetics, while R. H. gave a considerable amount in such training.

The distribution of Gross Phonetic Experience Scores.—Table 5 presents the distribution of Gross Phonetic Experience Scores for the 230 pupils of Grade III A. Since each Gross Phonetic Experience Score is the sum of the scores of a pupil's teachers for each half grade, it is, in the case of the III-A pupil, the sum of five scores made by his teachers on the Teachers' Blank. The lowest possible score a teacher could make was 25. Thus the lowest possible score for five teachers is 125. Likewise, since the highest possible score on the Teachers' Blank is 100, the highest possible Gross Phonetic Experience Score is 500. The possible range of Gross Phonetic Experience Scores is, therefore, from 125 to 500. The range of scores for 230 pupils was found to be from 160 to 349. As was to be expected from the scores on the Teachers' Blanks, the upper ranges of the distribution which represent extremely large amounts of

phonetic experience contain no scores. Nevertheless, the distribution indicates a considerable variation in amounts of phonetic experience.

C. THE TESTING PROGRAM

The group testing.—The group testing occupied the first few days of the testing program. It consisted in giving the *Otis Intelligence Scale,* Primary Examination: Form A;[3] the *Gates Silent Reading Test:* Types A, B, C, and D;[4] and the *Pressey Diagnostic Test:* Vocabulary—Grades 1 A-3 A.[5] The tests were administered under

TABLE 5
DISTRIBUTION OF GROSS PHONETIC EXPERIENCE SCORES
(III-A PUPILS IN RALEIGH)

Score	Frequency
340-349	10
330-339	0
320-329	9
310-319	18
300-309	15
290-299	21
280-289	20
270-279	15
260-269	26
250-259	15
240-249	2
230-239	3
220-229	9
210-219	13
200-209	4
190-199	15
180-189	2
170-179	22
160-169	11
Total.....	230
Md.....	267

standard conditions by Dr. W. A. Brownell, the writer, and graduate and senior members of classes in Experimental Education and Educational Measurements of Duke University. The members of these classes had been given full instruction with respect to the administration of the tests.

The individual testing.—Pupils were then tested with a battery of individual tests consisting of four types: (a) tests of phonetic

[3] Yonkers-on Hudson, New York: World Book Company.
[4] New York: Bureau of Publications, Teachers College, Columbia University.
[5] Bloomington, Illinois: Public School Publishing Company.

ability, (b) a test of word pronunciation ability, (c) tests of oral reading, and (d) a test of eye-voice span.[6]

The tests of phonetic ability comprised four of the *Tests for Phonic Abilities* (or extensions of these tests) devised by Arthur I. Gates.[7] Test A4 is a test of the ability to translate printed phonograms into sounds. Test A5 is similar except that it combines two phonograms. The tests are described by Gates as "visual stimulus" tests.[8] Tests B2 and B3, on the other hand, are tests of responses to auditory stimuli. They yield measures of the ability to give letter equivalents of sounds, B2 of single syllables, and B3 of combinations of two syllables.[9]

The test of word pronunciation used was the *Gates Graded Word Pronunciation Test:* Form II.[10] In this test, the child pronounces as many as possible of a list of 100 words of increasing difficulty.

The *Gray Oral Check Tests:* Sets II and III[11] were used to measure oral reading abilities. The tests provide a useful analysis of errors. The pupil reads a paragraph aloud, and errors of various types are recorded. Set II was constructed for Grades II and III, and Set III, for Grades IV and V. Each of these sets contains three paragraphs of fifty words. Both sets were used in order to measure both the upper and lower ranges of reading ability. The paragraphs in Set III were given in order to make sure that the pupils faced material representative of a new reading situation. In this study, Set II and two paragraphs of Set III were given.

A test patterned after the Buswell Eye-Voice Span Test[12] was given in an attempt to determine whether or not phonetic training

[6] Detailed descriptions and methods of administering these tests are given in: Agnew (2), Chapters VI, VII, and VIII. Samples of the tests together with keys for scoring are given in the same place, Appendix B.

[7] Gates (16), pp. 380-388.

[8] In order to make Test A6 more reliable, the number of items was increased from seven to twenty.

[9] Test B2 was modified by adding ten comparable items. This made twenty items in all. Test B3 was increased from eight to fifteen items.

The reliabilities of the four tests were calculated by the split-half method and the Spearman-Brown Prophecy Formula. The reliability coefficients (100 cases) are presented in the following table:

Test	Coefficient of Reliability	Test	Coefficient of Reliability
A4	.88	B2	.91
A5	.88	B3	.89

[10] New York: Bureau of Publications, Teachers College, Columbia University.

[11] Bloomington, Illinois: Public School Publishing Company.

[12] Buswell (5), pp. 87-88.

shortens the eye-voice span as is sometimes claimed. This test presents a little story to be read aloud, in which words of the same spelling occur as words of different meaning and pronunciation, e.g., "wind" from the verb "to wind," and "wind" as in "the wind blows." In order to pronounce these words correctly, it is necessary to allow the eyes to precede the voice so that the context may furnish a cue to the correct pronunciation.

In all, eight different individual tests were given to more than 300 third-grade pupils. This task was accomplished during the two weeks following the group testing. The testing was done by the corps of testers mentioned above as having administered the group tests. Care was taken to keep the conditions of testing uniform. Children were given the tests in several sittings so that fatigue was kept at a minimum.

CHAPTER IV

THE METHODS OF TREATING THE RESULTS

Since the methods of treating the results were somewhat complex, it is necessary to devote a chapter to explanation. The purpose was to divide the cases into groups representing different amounts of phonetic training and to compare these groups in terms of the tests of reading abilities. Two general methods were used for this purpose: (a) comparison of groups based on the Gross Phonetic Experience Scores, and (b) comparisons of patterns of phonetic experience, the different patterns representing different amounts of training at different times in the pupils' school experience.

A. COMPARISON BASED ON THE GROSS PHONETIC EXPERIENCE SCORES

The comparison of the scores on the tests of reading ability between pupils who had experienced large amounts of phonetic training and pupils who had experienced little phonetic training was made possible by a process involving three steps.

(1) The first step was to select groups at the extremes of the distribution of the Gross Phonetic Experience Scores.[1] In order to do this, arbitrary limits were set for the extremes. Thus, pupils with Gross Phonetic Experience Scores below 230 were included in the low group and pupils with Gross Phonetic Experience Scores above 290 were included in the high group. This method yielded 89 pupils in the low group, 86 pupils in the high group, and omitted the 55 in the middle of the distribution.

(2) The next step was to equate these extreme groups on the basis of measures of intelligence. This was done by pairing pupils in terms of M.A. and I.Q. Cases that could not be suitably paired were disregarded. By this means, two distributions were obtained, each containing 43 individual scores. The distributions of M.A.'s and I.Q.'s are given in Table 6. Hereafter, in this report, the group with the high Gross Phonetic Experience Scores is called Group G_H, and the group with the low Gross Phonetic Experience Scores is called group G_L. The G's indicate that the groups are based on the Gross Phonetic Experience Scores, and the subscripts indicate high and low amounts of phonetic experience.

[1] See Table 5.

TABLE 6

DISTRIBUTIONS OF M.A.'S AND I.Q.'S FOR GROUPS G_H AND G_L*

M.A. in Terms of Months	Frequency		I.Q.	Frequency	
	Group G_H	Group G_L		Group G_H	Group G_L
130-134	1	1			
125-129	1	1	125-129	1	1
120-124	8	8	120-124	3	5
115-119	8	8	115-119	4	5
110-114	12	12	110-114	12	14
105-109	4	4	105-109	6	4
100-104	3	3	100-104	5	3
95-99	4	4	95-99	7	5
90-94	0	0	90-94	3	4
85-89	1	1	85-89	0	0
80-84	1	1	80-84	2	2
Mean	112.04	112.04		107.50	108.08
σ	10.00	10.00		10.15	10.20

*Group G_H is made up of pupils with high gross Phonetic Experience Scores, and G_L, of pupils with low gross Phonetic Experience Scores.

(3) The scores of these groups on the tests of reading ability were compared in terms of means, P.E.'s of means, and critical ratios. The results are presented in Table 10 (Chapter V).

Limitation of the method of comparison.—The gross scores give equal value to phonetic experiences at different grade levels. This effect may not be what is wanted; possibly the method may yield measures so coarse that they obscure real and important differences. There may be, for example, a critical point in the child's educational experience that is particularly propitious for effective phonetic instruction. A child who had very little phonetic instruction in Grade II, but large amounts of phonetic experience in Grades I and III might be placed in the phonetic group. But if he had had little experience with phonetics in Grade II, the critical time for such instruction, this fact would be obscured if the gross scores alone were considered. In order to bring out the effects of phonetic instruction at different grade levels, another method of treating the data was necessary. This method is outlined in the following paragraphs.

B. COMPARISONS OF PATTERNS OF PHONETIC EXPERIENCE

Method used to isolate the factor of the time at which pupils obtained phonetic experience.—As has been brought out in the previous discussion, the object of the investigation was to isolate not

only the factor of gross amounts of phonetic experience, but also the factor of the time at which this experience had been obtained. In these comparisons, amounts of phonetic experience at different levels had to be identified. A method involving three steps was devised to accomplish these ends. According to this method, (1) each grade was divided into "phonetic," "medium," and "nonphonetic" groups; (2) the "medium" group was omitted; and (3) the "phonetic" and the "nonphonetic" groups were equated with respect to the sum of the phonetic experience scores for the other two grades. The resultant groups represented a pair of patterns of phonetic experience. In order to illustrate this process, the derivation of one pair of such patterns will be described in detail.[2]

The derivation of patterns A_p *and* A_n.—In order to measure the effects of varying amounts of phonetic experience in Grade I, the amounts in Grades II and III had to remain constant. There were three major steps in the process of deriving the pair of patterns representing the phonetic and nonphonetic groups in Grade I. (1) The first step was to make a distribution of the phonetic experience scores of pupils in Grade I as shown in Table 7.

(2) Since only the extremes are important for purposes of comparison, the second step was to choose arbitrary limits for the phonetic and nonphonetic groups and to disregard a group in the middle part of the distribution. Eighty cases with scores above 120 were selected to make up the phonetic group, and 111 cases with scores below 90, the nonphonetic group. The 39 cases between 90 and 119 were disregarded. At this stage no attention was paid to amounts of phonetic training above the first grade.

(3) The third step requires considerable explanation. The division into phonetic and nonphonetic groups (step 2) was made purely on the basis of the Phonetic Experience Scores for Grade I.

[2] Three other methods were thought of in connection with the isolation of these variables, but these methods were found to be impractical. (1) One method of isolating the factors, amount and time, is to consider the variable amount, in terms of large, medium, and small amounts, and to consider the variable time in terms of the three grade divisions. If these variables were broken down into their possible combinations it would yield twenty-seven groups. If there were 230 cases in all, the average number of cases in each group would be about eight. It is obvious that a comparison of such small groups would not yield reliable results. (2) Another method might be to consider the twenty-seven categories mentioned above as an unordered series and correlate it with the scores on the tests by means of the formula for η. [See Holzinger (25), pp. 266-277.] Interpretation of the results obtained by this method was found to be very difficult. (3) The method of partial correlation, in which amounts of phonetic experience are partialed out for each grade, was not used because the distributions of variables were by no means normal. [See Agnew (2), pp. 69-70.]

TABLE 7

FREQUENCY DISTRIBUTION OF PHONETIC EXPERIENCE SCORES OF PUPILS IN GRADE I

Scores	Frequency
130-139	17
120-129	63
110-119	28
100-109	9
90-99	2
80-89	4
70-79	65
60-69	40
50-59	2
Total	230

In order to keep constant the amounts of phonetic training in Grades II and III, the two groups, phonetic and nonphonetic, were equated in terms of their phonetic experience scores in these two grades. Two distributions were made of the scores for Grades II and III combined, one representing the phonetic group in Grade I and the other the nonphonetic group in Grade I. These distributions are shown as "initial distributions" in Table 8. Two new distributions were then obtained by elimination from one group or the other until

TABLE 8

DISTRIBUTIONS OF PHONETIC EXPERIENCE SCORES FOR GRADES II AND III COMBINED BEFORE AND AFTER EQUATING

Scores	Initial distributions		Equated distributions	
	A_p	A_n	A_p	A_n
200-209	14	8	8	8
190-199	15	1	4	1
180-189	6	1	6	1
170-179	2	23	2	22
160-169	14	3	11	2
150-159	14	9	12	7
140-149	2	0	1	0
130-139	1	2	1	2
120-129	11	13	7	9
110-119	0	15	0	0
100-109	1	24	0	0
90-99	0	11	0	0
80-89	0	1	0	0
N	80	111	52	52
Mean			167.1	166.9
σ			26.1	25.5

nearly equivalent means and standard deviations were secured. The final distributions are also shown in Table 8 under the heading "equated distributions."

For the sake of ease of expression and reference, the phonetic group in Grade I was called pattern A_p and the nonphonetic group A_n, the subscripts being abbreviations of "phonetic" and "nonphonetic." Thus pattern A_p means that the group of children so identified had had a larger amount of phonetic instruction in Grade I than had A_n, while both groups had obtained an equal amount of training in Grades II and III.

Other patterns.—In all, six pairs of patterns were isolated in this manner. In each case, procedures corresponding to those described in respect to the derivation of patterns A_p and A_n were followed. An analysis of Table 9 indicates in what grades the phonetic experience was varied and in what grades it was kept constant in the different patterns.

TABLE 9

DISTRIBUTIONS OF EQUATED GROUPS*

Pattern	Limits	Grades in which phonetics varied	Constant grades (amount of phonetics equated)	N	Mean	Standard Deviation
A_n.....	90 and below	I	II and III	52	167.1	26.1
A_p.....	120 and above	I	II and III	52	166.9	25.5
B_n.....	90 and below	II	I and III	50	136.2	22.1
B_p.....	120 and above	II	I and III	50	136.2	21.2
C_n.....	42 and below	III	I and II	45	228.3	18.7
C_p.....	50 and above	III	I and II	45	229.3	20.0
D_n.....	130 and below	II and III	I	56	90.5	20.8
D_p.....	165 and above	II and III	I	56	90.4	20.5
E_n.....	129 and below	I and III	II	53	107.9	26.3
E_p.....	165 and above	I and III	II	53	108.4	27.3
F_n.....	195 and below	I and II	III	47	40.7	7.6
F_p.....	230 and above	I and II	III	47	40.9	7.4

*Table 9 may be read as follows: Pattern A_n consists of all the cases having a phonetic score below 90 in the first grade that could be equated with cases in pattern A_p which consists of those who had a phonetics score in the first grade of more than 120. Equating was done on the basis of the sum of the phonetics scores for Grades II and III. The number of cases in A_n (and A_p) is 52, the mean is 167.1, and the standard deviation is 26.1.

Degree to which equating was possible.—Table 9 shows the degree to which equating was possible. For instance, in the case of patterns A_p and A_n, there is only .2 difference between the means of the two distributions, and there is only .6 difference between the standard deviations. For patterns B_p and B_n, the difference between the

means is .0 and that between the standard deviations .9. For patterns C_p and C_n, the difference between the means is 1.0 and between the standard deviations 1.3. Thus, it is evident that the sets of scores for each pair of patterns were equated closely.

Defense of the derivation of patterns.—It should be pointed out in this connection that the groups so equated were relatively small. However, it is to be borne in mind that the small size of the groups is more apparent than real. The description of the method employed in selecting these cases has emphasized the similarity of the subjects with respect to chronological age, level in school achievement, and similarity of instruction. That is, the subjects employed possess a really unusual degree of homogeneity. When groups of from 45 to 56 are compared in this study, therefore, they represent, in reality, much larger populations.

Comparison of groups and patterns in terms of the scores on the tests of reading ability.—The groups G_H and G_N were compared by calculating the differences between the means of the test scores of reading abilities. Since the groups had been equated in terms of intelligence, the following formula was used:[3]

$$\text{P.E.}_{\text{diff.}} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\left(1-r^2\right)} \cdot .674$$

Differences between each pair of patterns (in terms of the difference between the means of the scores of the various tests) were similarly calculated except that in these cases the usual formula for the critical ratio was used.

[3] See E. F. Lindquist, "The Significance of a Difference Between Matched Groups," *Journal of Educational Psychology*, XXII (March, 1931), 197-204.

CHAPTER V

RESULTS AND CONCLUSIONS OF THE RALEIGH INVESTIGATION

The results of the Raleigh investigation are presented in Table 10.[1] Before these results are analyzed, it is thought wise to consider certain factors that might have influenced the test scores.

TABLE 10

SUMMARY OF THE DIFFERENCES IN TERMS OF THE P. E. OF THE DIFFERENCES*

Tests	Groups and Patterns						
	G	A (I)	B (II)	C (III)	D (II-III)	E (I-III)	F (I-II)**
Gates A4	1.66	.42	2.15	.71	—1.68	4.56	—2.53
Gates A5	— .84	1.12	.52	1.36	—1.15	.34	—3.69
Gates B2	1.22	— .21	.34	—1.25	— .96	1.88	—3.97
Gates B3	3.02	1.27	.75	—1.88	— .26	6.41	—5.74
Gates Word Pronunciation	—2.40	— .85	1.21	1.79	2.30	1.98	—3.72
Gates,							
Type A	—2.50	1.68	— .89	1.22	—1.82	1.18	—8.49
Type B	—2.73	— .04	— .75	.15	—3.17	1.98	—5.72
Type C	—1.92	—1.40	— .61	.16	— .49	2.84	—6.35
Type D	— .87	—1.23	.39	— .75	—1.35	—2.83	—6.63
Pressey Vocabulary	.12	— .26	— .90	2.03	—1.33	—8.25	3.73
Gray II (errors)	.30	—1.16	1.02	— .91	.15	1.32	—3.54
Gray III (errors)	1.43	—1.16	2.10	.48	1.40	2.55	—5.13
Gray II (time)	— .32	— .80	2.46	— .00	1.60	1.50	—3.20
Gray III (time)	.85	— .33	2.46	.02	2.11	4.52	—5.17
Eye-voice Span	1.68	1.22	.89	1.95	— .09	3.54	—3.53
Mean	— .08	.00	.65	.33	.18	1.56	—4.17

*Negative differences favor the nonphonetic groups.
**Numbers in this row represent grades of variation.

Factors other than phonetic experience that might influence the test results.—A list of the chief factors other than phonetic training that might influence the test scores is given below:

1. Lack of homogeneity of subjects in age, school experience, etc.

[1] A complete exposition of the basic data (except those on word pronunciation) including the distributions of scores, the actual means and standard deviations, is given in: Agnew (2), Chapters VI, VII, and VIII. A similar exposition of the data on word pronunciation is given in: Agnew (1), Chapters III, IV, and V.

2. Biased sampling of schools and teachers.
3. Differences in the intelligence of the groups compared.
4. Differences in methods of teaching.
5. Differences in instructional material.
6. Differences in the motivation of reading.
7. Other unknown factors.

The factors listed above are briefly discussed in the following paragraphs.

1. The previous discussion has emphasized the homogeneity of the subjects.[2] It will be recalled that all the pupils selected had made normal progress in school, were in the same grade, and lived in the same city where they had received all of their school experience.

2. The pupils used in the study came from nine different schools (all the schools for white children in the city). It is interesting to point out in this connection that each pattern or group was composed of pupils from at least four different schools, and in many cases, from as many as six or seven schools. This fact insures a wide sampling of teachers, so that the chances are against the differences being due to the peculiar excellences or deficiencies of the teachers.

3. Table 11 presents the mean mental ages of the pupils represented by the patterns, together with the standard deviations of the means. The largest difference is in the case of the difference between the patterns B_p and B_n. This represents an apparent difference of five months. The difference is just great enough to be reliable, that is, it is assured that it is greater than 0. The correlation

TABLE 11

MEANS AND STANDARD DEVIATIONS OF THE M.A.'S FOR THE VARIOUS PATTERNS

Pattern	Mean	Standard deviation
A_p	108.0	10.15
A_n	111.9	10.20
B_p	115.0	9.85
B_n	110.0	11.95
C_p	111.6	9.05
C_n	115.6	10.85
D_p	115.5	8.45
D_n	111.8	10.30
E_p	112.8	8.75
E_n	115.5	11.55
F_p	117.9	10.05
F_n	117.8	10.50

[2] See Chapter III.

between intelligence and the test scores would have to be rather high for this small difference to play an important role in causing differences in the test scores. There is some evidence that this correlation is low. Table 12 shows the coefficients of correlation between the test scores and the intelligence of the matched groups G_H and G_L. Since these coefficients were really obtained by correlating the test scores of the two groups matched for intelligence, the true correla-

TABLE 12

CORRELATIONS BETWEEN TEST SCORES OF GROUPS G_H AND G_L (GROUPS MATCHED FOR M.A. AND I.Q.)

Tests	r	Tests	r
Gates A4	.06	Gates Types A, B, C, and D	.16
Gates A5	—.11	Pressey Vocabulary	—.08
Gates B2	—.06	Gray II, errors	—.13
Gates B3	—.21	Gray III, errors	—.07
Gates Type A	.06	Gray II, time	.09
Gates Type B	.20	Gray III, time	.07
Gates Type C	.30	Eye-voice Span	.20
Gates Type D	—.02		

tions between the scores and intelligence may be somewhat larger. It is doubtful, however, that the correlations would be large enough to modify appreciably the reliability of the differences.

It is not proposed to rule out intelligence as a possible factor in influencing test scores, but it is suggested that it is unlikely that the small differences in intelligence found between the patterns could have been responsible, to any great extent, for the differences between the means of the measures of reading ability.

4. There is a lack of information with regard to the methods of teaching reading, other than the degree to which it was phonetic or nonphonetic. The methods, however, were under the same supervision throughout the schools; and there is no reason to believe that the methods were causative in producing the differences in the means of the reading tests.

5. Since all the schools were in the same system, there is no reason to believe that the instructional materials varied significantly, or if so, that they tended to favor one pattern or group more than another.

6. Nothing is known about differences in motivation or other factors.

7. It is assumed that "unknown factors" operate equally upon all groups compared.

The influence of phonetic experience on phonetic abilities.—The tests of phonetic abilities, Gates A4, A5, B2, and B3, have been shown to have high reliability coefficients (see Chapter III, note 9). They test phonetic abilities as these are measured by the spelling and pronunciation of nonsense letter groups. The differences that are statistically reliable, as shown in Table 10, are consistent for each pair of patterns, as far as the four tests are concerned, but inconsistent as between patterns. Thus, the two reliable differences between the E-patterns favor the phonetic group. On the other hand, the differences between the F-patterns, although similarly consistent with each other, favor the nonphonetic group. If the E-patterns are taken alone, one would be led to conclude that phonetic training in Grades I and III results in higher scores on the Gates Tests A4 and B3. If the F-patterns are considered alone, it might be concluded that, when little phonetic training is given in Grades I and II, higher scores are obtained on the Gates Test A5, B2, B3, and possibly A4. In the latter instance it might have been concluded that pupils learned phonetic abilities without having had formal training in phonetics. While these facts may actually represent the situation, the apparent inconsistency between the results of the E- and F-patterns casts doubt on the significance of the differences obtained.

The influence of phonetic training on word pronunciation.—The comparison of G_p with G_n on the Gates Word Pronunciation Test reveals a small difference in favor of the nonphonetic group. A comparison of the patterns reveals: (a) phonetic training in Grade I seems to have a slight detrimental effect on word pronunciation ability; (b) phonetic training in Grade II seems to result in a slight increase of ability to pronounce words; (c) phonetic training in Grade III has a slightly greater tendency to increase ability to pronounce words; (d) the differences show a general lack of reliability.

The influence of phonetic training on silent reading abilities.—In the comparison of the patterns in terms of the results from the Gates Silent Reading Tests, Types A, B, C and D, five differences appear that are greater than three times the P. E. of the difference. All of these favor the nonphonetic groups, one favoring pattern D_n, and four, pattern F_n. The difference in the case of the D-patterns (for Type B) may not be significant because it is not supported by the other measures of silent reading ability (Types A, C, and D), but the large and consistent differences favoring the nonphonetic group in the comparisons of the F-patterns appear to be more significant. If this is the case, it may be concluded that there is some

evidence that large amounts of phonetic training in Grades I and II are not so advantageous to silent reading abilities (as measured by these tests) as are small amounts of phonetic training in these grades.

The influence of phonetic training on vocabulary.—In the comparisons of the patterns on vocabulary attainment, as measured by the Pressey test, one large difference appears to favor the nonphonetic group in the E-patterns, and one fairly large difference appears to favor the phonetic group in the F-patterns. It is possible that the phonetic group in pattern E_p was so accustomed to the phonetic attack on unfamiliar words that much time was taken in analyzing the nonsense words of the Pressey test. If this is true, the test may not have measured the actual vocabulary of the pupils in this group. However, if this were the case, there should be some evidence of this phenomenon in the comparisons of the C-patterns in which phonetic training in the third grade is isolated. Unreliable differences between the C-patterns favor the phonetic group. The C-patterns and the F-patterns are thus seen to be inconsistent with the E-patterns. This inconsistency tends to reflect doubt on the significance of the differences obtained.

The influence of phonetic training on oral reading.—In the comparisons of the patterns with respect to speed and accuracy on the Gray Oral Reading Check Tests (Gray II and III), there are five reliable differences. One, in the E-patterns, favors the phonetic group in the speed of reading Set III. The fact that the other differences between the E-patterns on both speed and accuracy are unreliable casts doubt on the significance of the one difference found. Reliable differences are found to favor F_n as opposed to F_p in both the measures of speed and accuracy. These differences alone would indicate a superiority on the part of nonphonetic training in Grades I and II in speed and accuracy on the Gray tests. However, no other comparisons of patterns bear out these conclusions, and again the differences seem to be of doubtful significance.

The influence of phonetic training on eye-voice span.—The test of eye-voice span is of questionable validity because it was necessarily short and seemed, in many cases, to involve too difficult reading material. On the other hand, although the instrument was somewhat crude, it was hoped that it might show extreme differences in eye-voice span (if they existed). The results, in no case, present differences as great as four times the P. E. of the difference.[3] Small

[3] The difference must be at least four times the P. E. of the difference in order to insure complete reliability. See Garrett (14), p. 136.

differences appear within the E-patterns and within the F-patterns. These differences are in opposite directions so that they are of questionable significance.

The influence of phonetic training on the battery of tests as a whole.—Table 10 presents (in the last row) an average of all the differences. This is, of course, a crude method of summary because it is impossible to say how the tests should be weighted. Also, any isolated reliable differences are obscured by being averaged with differences of small reliability. Yet these averages may aid in bringing out some general characteristics of the table as a whole.

1. Perhaps the most striking feature of these averages, and, in fact, of the comparisons as a whole, is the paucity of reliable differences. (a) Not a single difference as great as four times the P. E. of the difference appears in the comparison of Groups G_H and G_L. Only one difference (Gates B3) is as great as three times the P. E. of the difference. (b) In the comparisons of the phonetic and nonphonetic groups in the patterns A, B, and C (in which phonetic experience varied in one grade), only five differences are as great as three times the P. E. of the difference. (c) In the D-patterns (in which phonetic training varied in Grades II and III), only one difference is as great as three times the P. E. of the difference. (d) In the E-patterns (in which phonetic training varied in Grades I and III), five differences out of the fourteen are as great as three times the P. E. of the difference; and, of those, four favor the phonetic group, and one favors the nonphonetic group. Thus, in the above comparison there is no consistent evidence that the differences in phonetic training measured in the Raleigh study affected the test scores appreciably.

2. Only in the case of the comparisons of the F-patterns (in which phonetic training varied in Grades I and II), do differences appear consistently reliable. (a) With but one exception, the differences are as great as three times the P. E. of the difference. (b) All but one of these differences (that in vocabulary scores) favors the nonphonetic groups. This is true even in the phonetic tests. (c) The average of these differences is 4.17 P. E. It would seem that the comparison of the F-pattern in terms of the battery of reading tests presents rather consistent evidence that phonetic experience (in the first two grades) is not so beneficial to reading abilities, as measured by the tests, as is nonphonetic experience. How great this ben-

efit of nonphonetic training may be, however, is not answered from these data.[4]

3. The inner consistency of the differences found in the E-patterns suggests that these differences have some significance. If the difference favoring the nonphonetic group on the Pressey Vocabulary Test is omitted and the average of the other differences computed, the resultant mean is found to be 2.35. This difference seems to indicate a fairly adequate advantage on the part of the phonetic group.

Various interpretations that may be made of the inconsistencies between the directions of the differences in the E- and F-patterns are mentioned in the last paragraphs of this chapter.

General conclusions of the Raleigh investigation.—The conclusions suggested as a result of the Raleigh investigation may be summarized as follows:

1. The comparisons made failed to reveal a significant advantage or disadvantage (in terms of reading test scores) arising from different amounts of phonetic experience as measured by the Gross Phonetic Experience Scores.

2. The effort to find a critical grade in which phonetic experience is particularly effective for training in reading was unsuccessful.

3. There seems to be a tendency for large amounts of phonetic experience in Grades I and II (as is indicated in the F-patterns) to affect the reading abilities adversely.

[4] A search was made for factors that might account for the differences between patterns F_p and F_n.

(a) It was observed, as is indicated in Table 11, that the average M. A.'s of these two patterns were practically identical. The same is true of the standard deviations. The range of M. A.'s in both patterns was found to be approximately from 90 to 140. Both groups were somewhat above the average in M. A. It is difficult to say what significance (if any) this fact may have. It is possible that more intelligent children tend to be inhibited by phonetic training.

(b) It was thought that one or both of these groups might happen to be composed of pupils who had been in particular schools or who had had particular teachers. No facts were disclosed to indicate that any other factor than phonetic experience operated in the selection of groups. The table below shows the distribution of the groups of schools.

Schools	Frequencies		Schools	Frequencies	
	F_p	F_n		F_p	F_n
Hayes Barton	19	16	Thompson	0	5
Boyland Heights	11	4	Lewis	2	0
Olds	0	16	Pool	1	0
Murphy	8	6	Wiley	6	0
			Total	47	47

4. The direction of the differences found in the E-patterns suggests that large amounts of phonetic experience in Grades I and III are beneficial to most of the reading abilities measured.

Interpretation.—These conclusions may have a number of interpretations, the most reasonable of which appear to be the following:

1. The measures of phonetic experience may have been faulty.

2. In general, the differences between the amounts of phonetic experience as between groups or between members of patterns may not have been great enough to affect measurable differences in the reading test scores.

3. The inconsistency between the E- and F-patterns may be due to unknown factors in selection or training. It is perfectly possible, however, that the inconsistency is due to the actual differences in the patterning of the phonetic training. The differences in scores may be the effect of the interference of one type of training with another. Thus, phonetic training in the first two grades may be ineffective if, in the third grade, other types of training are stressed. The training in the third grade (immediately before the testing) might tend to interfere with the earlier training and thus cause the relatively low scores observed in pattern F_p. On the other hand, since pattern F_n represents a fairly consistent nonphonetic training, the interference would be less and the scores higher, as was actually the case.

In pattern E_p, the interference effect would not be so apparent, since the phonetic training in the third grade would tend to counteract any interference effect that might have occurred in the second grade. If this is true, the data suggest that consistent phonetic training might have beneficial effects on the abilities measured.

In view of the high reliability and the facts presented previously concerning the validity of the Teachers' Blank, it seems reasonable to assume that the measures of phonetic experience were not at fault. The second and third of the four explanations given above seem worthy of further investigation. The study to be reported in the subsequent chapters was undertaken to determine the effects of considerably larger and more consistent amounts of phonetic experience on the scores of the same battery of tests. In this way, it was anticipated that the explanations of the results of the Raleigh investigation might be checked.

CHAPTER VI

THE PURPOSE AND TECHNIQUE OF THE DURHAM INVESTIGATION

A. PURPOSE

The second investigation was undertaken (1) in order to check the results obtained in the Raleigh investigation, and (2), in order to provide new data on the effects of larger and more consistent amounts of phonetic experience than those found in Raleigh.

It will be recalled that the upper ranges of the possible scores on the Teachers' Blanks were not sampled in Raleigh. Table 3 presents the distribution of teachers' scores in Raleigh and shows the highest scores for the various grades to have been: for Grade I, between 65 and 69; and for Grades II and III, between 69 and 75. Since the highest possible scores on the Teachers' Blank is 100, a possible range of scores consisting of some 20 points at the upper range of the distribution was not sampled in Raleigh.

Furthermore, when the Gross Phonetic Experience Scores are considered (see Table 5), it will be observed that the upper limit of scores fell in the interval 340-349. The possible upper limit of these scores is 500. Thus, a possible range of 150 points was unsampled by the Raleigh data. A search was made, therefore, for third-grade pupils whose phonetic experience might be great enough to sample this upper range. It having been the policy during the preceding three years (1931-35) to teach large amounts of phonetics in Durham, that city was selected for the second investigation.

B. THE TECHNIQUE OF THE INVESTIGATION

Determination of amounts of phonetic experience.—Teachers' Blanks were submitted to the first-, second-, and third-grade teachers in a number of the Durham elementary schools. The blanks were scored as were those used in Raleigh.

Rough estimates of the pupils' phonetic experience were made by adding together the scores of trios of first-, second-, and third-grade teachers. After the schools had been selected (in which there seemed to be the greatest likelihood of high scores) accurate records were made of the pupils' phonetic experience, as was done in the Raleigh investigation.

Selection of schools.—Two schools were selected for investigation. In the selection of these schools, two principles were borne in mind: first, it was desirable to obtain subjects who had experienced large amounts of phonetic training; and second, it was desirable to obtain a distribution of subjects comparable to those used in the Raleigh investigation. The North Durham and Edgemont schools were finally selected as best meeting these criteria. Both of these schools had consistently emphasized phonetic instruction during the training of the pupils who were in the third grade at the time of the investigation. Neither school represents an extreme in the economic and cultural levels of the city. The Edgemont School derives a majority of its population from the homes of mill workers, people in small business concerns, etc. The population is probably somewhat below the average for the city in cultural advantages. The pupils of the North Durham School come, in general, from more advantageous home environments, perhaps slightly above the average for the city. Thus the schools furnished subjects not significantly different from, and certainly not superior to, the subjects of the Raleigh investigation.

Determination of amounts of phonetic training.—An analysis of the school records revealed the fact that, out of a third-grade population of about 200, there were 110 pupils who had made regular progress through the grades, and who had received all three years of their school training in these selected schools. Since there had been no changes in the teaching personnel during these years, and since Teachers' Blanks (covering the time when the pupils had been taught by the particular teachers) had been secured from all the first-, second-, and third-grade teachers, it was possible to compute the amounts of phonetic experience for each pupil in terms of a Gross Phonetic Experience Score.

Table 13 presents a frequency distribution of the Gross Phonetic Experience Scores of the 110 selected pupils in Durham. It is apparent that the pupils had received consistently large amounts of phonetic training. When it is recalled that, in Raleigh, no Gross Phonetic Experience Scores exceeded 350, and only a scattered few exceeded 320 (see Table 5), it is clear that these Durham pupils had received considerably more phonetic instruction than had the pupils in Raleigh.

Administration of the tests.—The same battery of tests that had been previously administered in Raleigh was given in the selected Durham schools. A class in Educational Measurements from Duke

TABLE 13

FREQUENCY DISTRIBUTION OF GROSS PHONETIC EXPERIENCE SCORES (110 CASES, DURHAM)

Score	Frequency
390-399	20
380-389	0
370-379	90
Total	110

University, composed of seniors and graduate students, administered the tests. The administrators were carefully instructed in the technique of giving the tests in order that the conditions of testing that had been set up in Raleigh might be duplicated as nearly as possible. The group tests, the intelligence test, the vocabulary test, and the tests of silent reading abilities were given to all the third-grade pupils. The individual tests of phonetic abilities, eye-voice span, oral reading and word pronunciation were given to the 110 selected pupils.

Treatment of the results.—It will be recalled that the primary purpose of the Durham investigation was to provide data on the reading abilities of pupils who had received large consistent amounts of phonetic training, in order that these data might be compared with

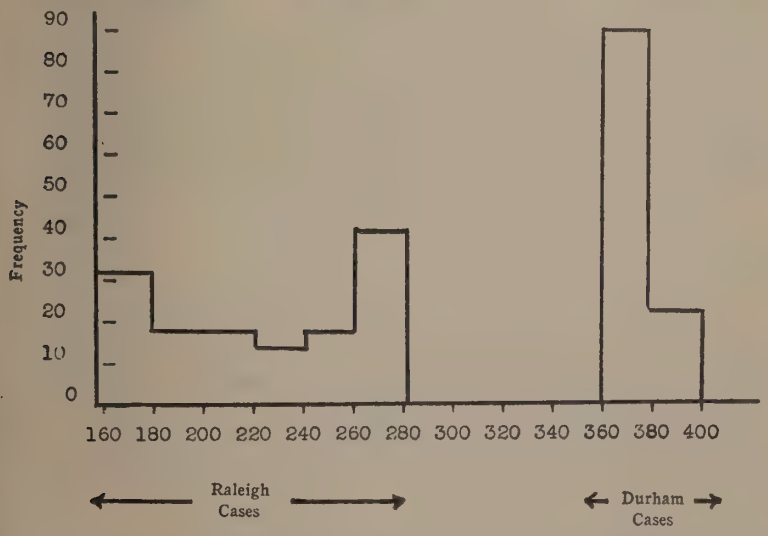

FIG. 1. Histogram showing the frequency of the Gross Phonetic Experience Scores from which groups were selected in the comparison of the Raleigh and Durham data.

the data obtained in Raleigh, where the amounts of phonetic training had been smaller and less consistent. In order to make the groups compared represent similar ranges of intelligence, pupils with low Phonetic Experience Scores in Raleigh were paired with Durham pupils (whose Phonetic Experience Scores were high) on the basis of M.A.'s and nearly equivalent I.Q.'s. A detailed account of this procedure follows:

Figure 1 indicates graphically the distribution of Phonetic Experience Scores from which the pairs were drawn.

All the Raleigh cases had Phonetic Experience Scores that fell in the range from 160 to 270, and all the selected Durham cases had Phonetic Experience Scores ranging from 370 to 400.[1]

By the method of pairing, 89 cases were selected from the Raleigh group, and a similar number from the Durham group. Table 14 presents the distributions of the M.A.'s and I.Q.'s of the two selected groups. Most pairs had identical M.A.'s, and in no case did the M.A.'s of a pair differ more than two months. Care was exercised

TABLE 14

FREQUENCY DISTRIBUTIONS OF M.A.'S AND I.Q.'S OF THE RALEIGH AND DURHAM GROUPS SELECTED BY PAIRING

M.A. in Terms of Months	Frequencies		I.Q.	Frequencies	
	Raleigh	Durham		Raleigh	Durham
140-149	1	1			
130-139	7	7	130-139	1	1
120-129	27	27	120-129	18	15
110-119	37	37	110-119	37	37
100-109	12	12	100-109	20	22
90-99	5	5	90-99	10	11
			80-89	2	2
			70-79	1	1
Total	89	89		89	89
Mean	117	117		112	111

[1] The two groups differ, not only in terms of gross amounts of phonetic experience, but in terms of the patterns of that experience. In the Raleigh group, the training, although fairly consistently nonphonetic, in some instances varied considerably in individual grades. Thus a pupil in Grade I might have a score of 50 or 140 and still be included in the group. Similar variations sometimes occurred in other grades. It is possible, therefore, that the interference factors, which may have operated to lower the reading scores in pattern F_p, operated in some cases to lower the reading scores of the Raleigh group.

The Durham group, on the other hand, represents highly consistent phonetic training for all three grades. Thus, in this group, interference was probably not a factor.

in order to avoid wide discrepancies between the I.Q.'s of the members of a pair. In no case did the I.Q.'s of a pair differ more than eight points.

In order to express simply the nature of the groups, Table 15 has been prepared. A glance at the table will reveal the constants and the variables in the known conditions of the two groups.

TABLE 15

FACTORS IN THE COMPOSITION OF THE DURHAM AND RALEIGH GROUPS

Conditions	Raleigh	Durham
Number of cases	89	89
Number of years in school	2.6-2.7	2.6-2.7
Intelligence	Equated by pairing with Durham	Equated by pairing with Raleigh
Time spent on reading (daily)*		
Grade I	40 to 70 minutes	60 to 100 minutes
Grade II	60 to 100 minutes	60 to 95 minutes
Grade III	60 to 80 minutes	60 to 80 minutes
Number schools represented	7	2
Number teachers represented	32	10
Class organization	Home Room	Platoon
Supervisor's attitude toward phonetics	Opposed to direct phonetic teaching	Favored considerable phonetic teaching
Range of Phonetic Experience Scores	160-270	360-400

**Time devoted to reading in Raleigh.*—The basis of the estimates of the time spent in reading in the Raleigh primary grades was an analysis of the sample daily programs given in Curriculum Bulletins issued by the Raleigh Public Schools. The teachers, whose samples are given, were all teachers who filled out blanks in the Raleigh investigation.

Time devoted to reading in Durham.—The estimates of time spent in reading activities in the Durham schools were obtained through the kindness of Mrs. Robinson, Supervisor of the Durham Elementary Schools, who analyzed the daily program for this purpose.

CHAPTER VII

RESULTS AND CONCLUSIONS OF THE DURHAM INVESTIGATION

Table 16 presents the results of the comparison of the Raleigh scores with those from the Durham schools. An analysis of the table reveals a remarkably consistent picture of superiority in the Durham scores except in certain cases in which silent reading and speed in oral reading were measured.

The phonetic tests.—The comparative scores of the four phonetic tests, Gates A4, A5, B2, and B3, indicate a definite superiority on the part of the Durham group. In every case the difference between the means is reliable. The differences vary from seven to more than twelve times the P.E. of the difference. Thus, in so far as they measure phonetic abilities, the tests indicate that the pupils in the Durham schools had developed phonetic skills distinctly more than had the Raleigh pupils.

The word pronunciation tests.—That these phonetic abilities carry over into the pronunciation abilities is borne out by the fact that the Durham pupils obtained much higher scores on the Gates Word Pronunciation Test than did the Raleigh pupils. Here the difference is seen to be more than eight times the P.E. of the difference.

The silent reading tests.—It will be recalled that the Gates Silent Reading Tests, Types A, B, C, and D, were given to test specific types of silent reading abilities. In Table 16 the scores of these tests have been reduced to grade equivalents. In Types A and B (Reading to Appreciate the General Significance, and Reading to Predict the Outcome of Given Events), no reliable difference appears between the two groups. In Type C (Reading to Understand Precise Directions), a small reliable difference favors the Durham group. A somewhat less reliable difference favors the Durham group in Type D (Reading to Note Details). The differences are the smallest that appear in the table. It will be observed that if the norms are reliable, the averages of both the Durham and Raleigh groups are slightly above the expected grade-equivalent of about 3.8.[1]

The vocabulary tests.—A difference of more than ten times the

[1] The norms used were taken from Arthur I. Gates's *Manual of Directions for Gates Silent Reading Tests* (revised January, 1934), Bureau of Publications, Teachers College, Columbia University.

TABLE 16

DIFFERENCES BETWEEN THE RALEIGH AND DURHAM GROUPS IN TERMS OF THE DIFFERENCES BETWEEN THE MEANS OF THE TEST SCORES (89 CASES IN EACH GROUP)

Test	Group*	Mean	Differences between Means**	P.E. of Differences***	Critical Ratio
Gates A4.....	D	79.50			
	R	63.31	16.19	1.61	10.05
Gates A5.....	D	32.17			
	R	23.85	8.32	1.17	7.11
Gates B2.....	D	29.29			
	R	18.11	11.18	.93	12.02
Gates B3.....	D	15.20			
	R	9.29	5.91	.70	8.44
Word Pronunciation .	D	70.17			
	R	53.15	17.02	1.92	8.86
Gates Type A.....	D	4.08			
	R	4.03	.5	.09	.55
Gates Type B.....	D	4.18			
	R	4.18	.00	.11	.00
Gates Type C.....	D	4.61			
	R	4.11	.50	.12	4.16
Gates Type D.....	D	4.38			
	R	4.15	.23	.08	2.87
Pressey Vocabulary..	D	71.85			
	R	59.26	12.57	1.21	10.39
Gray Set II.....	D	2.35			
(errors)	R	8.79	6.44	.76	8.47
Gray Set III.....	D	7.05			
(errors)	R	17.50	10.45	.83	12.54
Gray Set II.....	D	73.04			
(time)	R	38.78	—40.26	2.34	17.20
Gray Set III.....	D	77.48			
(time)	R	52.87	—26.61	3.09	8.61
Eye-Voice Span.....	D	37.94			
	R	31.69	6.25	.64	9.76

*D and R refer to Durham and Raleigh respectively.

**Negative differences favor the Raleigh group.

***The Lindquist formula for matched groups was used to determine the P.E. of the difference.

P.E. of the difference favors the Durham group on the Pressey Diagnostic Test of Vocabulary. The average score for the Raleigh pupils, 59, represents a vocabulary of 1,200 words; and the average score of the Durham pupils, 71, represents a vocabulary of 1,400 words.[2]

The oral reading tests.—The Gray Oral Check Tests yielded two types of measures, the number of errors and the time consumed in

[2] S. L. Pressey and L. C. Pressey, *Directions and Class Record Sheet for Pressey Diagnostic Reading Test.*

reading the passages. The results indicate that the Raleigh pupils made considerably more errors on both of the tests. The differences between the means of the Raleigh and Durham groups are statistically reliable, as is indicated by the critical ratios. On Set II the difference is more than eight times the P.E. of the difference; and on Set III, the more difficult of the tests, the difference is more than eleven times the P.E. of the difference. The differences are in the opposite direction in the case of the time taken in reading the passages. Differences, eight and nine times the P.E. of the differences, show that the Raleigh pupils read more rapidly than did the Durham pupils. Thus, the Durham pupils appear to be slower, but more accurate oral readers.

The eye-voice span test.—The averages of the eye-voice span test scores favor the Durham group by a difference of more than nine times the P.E. of the difference.

Speed and accuracy on the silent reading tests.—In the original study (Agnew 2, Chapter XI) a study was made of speed and accuracy on the silent reading tests. These data showed no consistent evidence that large amounts of phonetics made silent reading slower but more accurate, as appears to be the case in oral reading.

Methods used in word pronunciation.—An analysis of the methods used in pronouncing words in the Gates Pronunciation Tests showed that approximately 70 per cent of the Durham subjects used phonetic methods. On the other hand, it was found that only 30 per cent of the subjects in Raleigh used phonetic methods. These data indicate that the Durham subjects actually used phonetic methods in pronouncing words.

CONCLUSIONS

The conclusions of the Durham investigation may be summarized as follows:

1. The comparatively large and more consistent amounts of phonetic training received by the Durham pupils seem to have resulted in greater phonetic abilities as measured by the Gates phonetic tests.

2. The Durham pupils were superior to the Raleigh pupils in word pronunciation ability.

3. The study of methods used in word pronunciation on the Gates Graded Word Pronunciation Test revealed the fact that the Durham pupils used phonetic methods of word pronunciation to a much greater degree than did the Raleigh pupils.

4. Comparatively little difference appears between the Durham and Raleigh pupils in the silent reading abilities measured. Small differences on two of the four tests favored the Durham group.

5. No consistent differences appear between the two groups with respect to speed and accuracy on the silent reading tests.

6. The greater phonetic training of the Durham group seems to have resulted in the acquisition of greater vocabulary.

7. The Durham group appeared to be slower but more accurate on the oral reading tests.

8. The Durham pupils seem to have developed greater eye-voice span than the Raleigh pupils. This conclusion tends to refute the argument that phonetic training decreases the eye-voice span.

CHAPTER VIII

BRIEF SUMMARY OF THE RESULTS OF THE INVESTIGATIONS, AND CERTAIN EDUCATIONAL IMPLICATIONS

A. THE RELATION OF THE INVESTIGATIONS TO THE CONTROVERSIAL ISSUES WITH REGARD TO PHONETIC INSTRUCTION

In Chapter I, the arguments for and against phonetic training were summarized. The investigations reported in the present study present evidence that has direct bearing on a number of the arguments.

The investigations have tended to support four of the arguments in favor of phonetic training. These arguments are that phonetic training when given consistently in large amounts (as in Durham): (a) increases independence in recognizing words previously learned; (b) aids in "unlocking" new words by giving the pupil a method of sound analysis; (c) encourages correct pronunciation; and (d) improves the quality of oral reading. The investigations provided no evidence on the other arguments in favor of phonetic training.

The study tends to show that a number of the objections to phonetic training have been exaggerated. In other words, although the investigation offered opportunity for evidence in support of these objections, such evidence did not appear. There was no evidence that large consistent amounts of phonetic training tend: (a) to sacrifice interest in the content of reading; (b) to result in the neglect of context clues; (c) to result in unnecessarily laborious recognition of unfamiliar words; and (d) to be unnecessary because the advantages attributed to phonetic training might be obtained without formal training. Some positive evidence indicated too that (e) phonetic training does not narrow the eye-voice span.

On the other hand, there are some data to show that large amounts of phonetic training tend to slow up oral reading. This is, in a sense, counteracted by greater accuracy in oral reading.

The investigations did not reveal striking differences in silent reading ability as between groups having large differences in amounts of phonetic training. There was no evidence that phonetic training decreases efficiency in silent reading. This may be due to the fact that speed in silent reading is largely acquired in the grades above

the primary level. Further investigation would be necessary in order to determine the effects of this early training on silent reading in the advanced grades.

B. THEORETICAL CONSIDERATIONS

Interference as an explanation of the Raleigh and Durham results.—Numerous comparisons were made in the Raleigh investigation between phonetic and nonphonetic patterns of phonetic training. In these comparisons, with two noteworthy exceptions, no reliable differences were found. The two exceptions, those found in the E- and F-patterns, furnish clues to an explanation of the lack of reliable differences between the other pairs of patterns and the differences between the Raleigh and Durham results. The differences between the members of the E- and F-patterns may have been due to the factor of interference. If interference operated in these pairs of patterns unequally to produce the differences between the members of the patterns, it is possible that interference operated equally in the other pairs of patterns. This fact would tend to account for the lack of reliable differences between the members of the A-, B-, C-, and D-patterns.

Furthermore, the inconsistency of instruction represented in the Raleigh group may have been responsible, to an extent, for the lower reading scores of the Raleigh pupils as compared to the Durham pupils. Since the amount of phonetic training to which the Durham group had been subjected was consistently high for each grade, the factor of interference probably did not operate to lower the Durham scores.[1]

Phonetic abilities a function of factors other than amounts of training.—The foregoing considerations suggest that *phonetic abilities are not only a function of amounts of phonetic training, but also of the consistency of phonetic training.* If phonetic abilities were merely a function of amounts of training, it would be expected that the phonetic groups in the Raleigh investigation would manifest greater phonetic ability than did the nonphonetic groups. This is not the case, however. Although the phonetic groups in Raleigh received much more phonetic training than did the nonphonetic groups, the phonetic groups, in general, showed no superiority over

[1] An effort was made to study the effects of interference statistically by comparing the mean deviations from the means of the individual phonetic scores of the Raleigh pupils with the corresponding means of the Durham pupils. The comparison indicated that the Durham pupils' phonetic scores (for the various half-years) varied less from their means than was the case with the Raleigh pupils.

the nonphonetic groups in phonetic abilities. The situation is illustrated by the chart which follows (see Figure 2). In this chart, the dotted line represents the expected growth of phonetic ability if the amount of training were the only causative factor. The heavy line represents the relationship that seems actually to exist.

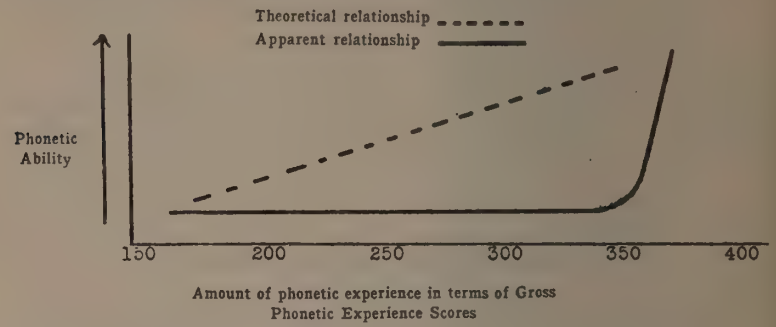

FIG. 2. Theoretical presentation of the Raleigh and Durham data.

The difference between these two lines may be theoretically explained by either or both of two hypotheses. First, it is possible that the factor of interference in the phonetic groups in Raleigh tended to keep the increasing amounts of phonetic training from increasing phonetic abilities. The facts to support this hypothesis have been presented in the foregoing paragraphs.

The second hypothesis (somewhat related to the first) is that phonetic abilities are highly complex, and that it is necessary, therefore, to have large amounts of practice in order to insure the acquisition of phonetic skills. Thus, *the learning of phonetic skills may be a function of the difficulty of those skills, as well as of amount of training.* That these skills, the ability to pronounce letter groups phonetically and to spell the sounds phonetically, are complex has long been recognized. That the many sounds of letters could be learned without a large amount of practice and drill seems improbable. Since these abilities are complex and require considerable practice to insure their use, it follows that, if phonetic methods are only partly learned, other methods of attack on words might be used in the testing situation; and the half-learned habits of phonetic analysis might serve rather to hinder successful responses than to improve those responses.

In the case of the Durham subjects, the methods of phonetics had been learned to the relative exclusion of other methods; in other

words, phonetics had become functional. This possibility is borne out by the very marked superiority of the Durham pupils on the Gates phonetic tests and by the common use of phonetic methods by the Durham pupils on the Gates Word Pronunciation Test.

C. EDUCATIONAL IMPLICATIONS

Should phonetic methods be employed in the teaching of primary reading? The answer to this question can be given only when the purposes of teaching primary reading have been agreed upon. If the basic purpose in the teaching of primary reading is the establishment of skills measured in this study (namely: independence in word recognition, ability to work out the sounds of new words, efficiency in word pronunciation, accuracy in oral reading, certain abilities in silent reading, and the ability to recognize a large vocabulary of written words), the investigations would support a policy of large amounts of phonetic training. If, on the other hand, the purposes of teaching primary reading are concerned with "joy in reading," "social experience," "the pursuit of interests," etc., the investigations reported offer no data as to the usefulness of phonetic training.

It is possible that the aims of primary reading should embrace all these purposes. If this is true, the relation of phonetic training (and the abilities resulting from phonetic training) to these other purposes would have to be determined before the place of phonetic training in primary reading instruction can be ascertained.

BIBLIOGRAPHY*

1. Agnew, Donald C. "A Study of Word Pronunciation by Third-Grade Pupils." Unpublished Master's Thesis, Duke University, 1932.
2. Agnew, Donald C. "The Effect of Varied Amounts of Phonetic Training on Primary Reading." Unpublished Doctor's Thesis, Duke University, 1936.
3. Atkins, Ruth E. "An Analysis of the Phonetic Elements in a Basal Reading Vocabulary," *Elementary School Journal,* XXVI (April, 1926), 596-606.
4. Bowden, Josephine H. "Learning to Read," *Elementary School Teacher,* XII (September, 1911), 21-33.
5. Buswell, Guy T. *An Experimental Study of Eye-Voice Span in Reading.* Supplementary Educational Monographs, No. 17. Chicago: Department of Education, University of Chicago, 1920. xii, 106 pp.
6. Buswell, Guy T. *Fundamental Reading Habits: A Study of Their Development.* Supplementary Educational Monographs, No. 21. Chicago: Department of Education, University of Chicago, 1922. xiv, 150 pp.
7. Cabell, Elvira D. "The Status of Phonics Today, Part I," *Chicago Schools Journal,* XIII (April, 1931), 370-373.
8. Cabell, Elvira D. "The Status of Phonics Today, Part III," *Chicago Schools Journal,* XIII (May, 1931), 430-434.
9. Carroll, H. A. "Effect of Intelligence upon Phonetic Generalization," *Journal of Applied Psychology,* XV (April, 1931), 168-181.
10. Cordts, Anna D. "An Analysis and Classification of the Sounds of English Words in a Primary Reading Vocabulary." Unpublished Doctor's Thesis, College of Education, State University of Iowa, Iowa City, Iowa, 1925.
11. Cordts, Anna D. "Facts for Teachers of Phonics," *Elementary English Review,* III (April, 1926), 116-121.
12. Currier, Lillian B. "Phonetics or no Phonetics," *Elementary School Journal,* XXII (February, 1923), 448-452.
13. Dreis, Theme A. "A Case of Remedial Reading," *Elementary School Journal,* XXXI (December, 1930), 292-300.
14. Garrett, Henry E. *Statistics in Psychology and Education.* New York: Longmans, Green and Company, 1926. xiii, 317 pp.
15. Garrison, S. C. and Heard, M. T. "An Experimental Study of the Value of Phonetics," *Peabody Journal of Education,* IX (July, 1931), 9-14.
16. Gates, Arthur I. *The Improvement of Reading.* New York: The Macmillan Company, 1927. xii, 440 pp.

* Since this monograph was written, the two following articles have appeared: Dolch, E. W., and Bloomster, Maurine, "Phonic Readiness," *Elementary School Journal,* XXXVIII (November, 1937), 201-205; and Tate, Harry L., "The Influence of Phonics on Silent Reading in Grade I," *Elementary School Journal,* XXXVIII (June, 1937), 752-763.

17. Gates, Arthur I. "Intrinsic Method in Reading," *Educational Outlook,* V (January, 1931), 105-114.
18. Gates, Arthur I. *Manual of Directions for Gates Silent Reading Tests.* Bureau of Publications, Teachers College, Columbia University. (Revised, January, 1934.)
19. Gates, Arthur I. *New Methods in Primary Reading.* Bureau of Publications, Teachers College, Columbia University, 1928. ix, 236 pp.
20. Gates, Arthur I. *A Reading Vocabulary for the Primary Grades.* Bureau of Publications, Teachers College, Columbia University, 1926. 23 pp.
21. Gates, Arthur I. *Psychology of Reading and Spelling with Special Reference to Disability.* Contributions to Educator, No. 129, Teachers College, Columbia University, New York, 1922. vii, 108 pp.
22. Gates, Arthur I. "Studies of Phonetic Training in Beginning Reading," *Journal of Educational Psychology,* XVIII (April, 1927), 217-226.
23. Gray, William S. *Directions for Giving and Scoring Standardized Oral Reading Check Tests.* Bloomington, Illinois: Public School Publishing Company.
24. Gray, William S. *Remedial Cases in Reading: Their Diagnosis and Treatment.* Supplementary Educational Monographs, No. 22. Chicago: Department of Education, University of Chicago, 1922. vii, 208 pp.
25. Holzinger, Karl J. *Statistical Methods for Students of Education.* New York: Ginn and Company. viii, 372 pp.
26. Horn, Ernest. "The Child's Early Experience with the Letter A," *Journal of Educational Psychology,* XX (March, 1929), 161-168.
27. Housh, E. T. "An Analysis of the Vocabulary of Ten Second-Year Readers," *Seventeenth Yearbook of the National Society for the Study of Education,* Part I. Bloomington, Illinois: Public School Publishing Company, 1918. 40-45.
28. Judd, Charles H. *Reading: Its Nature and Development.* Supplementary Educational Monographs, No. 4. Chicago: Department of Education, University of Chicago, 1918. xiv, 192 pp.
29. McKee, Paul. *Reading and Literature in the Elementary School.* Houghton Mifflin Company, 1934. x, 591 pp.
30. Morphet, Mabel V. and Washburne, Carleton. "When Should Children Begin to Read?" *Elementary School Journal,* XXXI (March, 1931), 496-503.
31. Mosher, Ramon M. "Some Results of Teaching Beginners by the Look and Say Method," *Journal of Educational Psychology,* XIX (March, 1928), 185-193.
32. Mosher, Ramon M. and Newhall, Sidney M. "Phonic Versus Look-and-Say Training in Beginning Reading," *Journal of Educational Psychology,* XXI (October, 1930), 500-506.
33. O'Brien, John A. *Reading: Its Psychology and Pedagogy.* New York: Century Company, 1926. xxviii, 308 pp.

34. Osburn, W. J. *Bulletin on Remedial and Follow-Up Work in Silent Reading,* No. 1. Bloomington, Illinois: Public School Publishing Company, 1925. 23 pp.
35. Packer, J. T. "The Vocabulary of Ten First Readers." *Twentieth Yearbook of the National Society for the Study of Education,* Part II. Bloomington, Illinois: Public School Publishing Company, 1921. 127-144.
36. Pressey, S. L. and Pressey, L. C. *Directions and Class Record Sheet for Pressey Diagnostic Reading Test.* Bloomington, Illinois: Public School Publishing Company.
37. Raleigh Public Schools. *Curriculum Bulletin No. 3. Teaching in the First Grade.* (1928, republished, 1931.)
38. Raleigh Public Schools. *Curriculum Bulletin No. 4. Teaching in Grades II and III.* (1928.)
39. Sexton, Elmer K. and Herron, John S. "The Newark Phonics Experiment," *Elementary School Journal,* XXVII (May, 1928), 690-701.
40. Smith, Nila B. *American Reading Instruction.* New York: Silver Burdett and Company, 1934. x, 287 pp.
41. Smith, Nila B. "The Construction of First Grade Reading Material," *Journal of Educational Research,* XVII (February, 1928), 79-89.
42. Tinker, Miles A. "Visual Apprehension and Perception in Reading," *Psychological Bulletin,* XXVI (April, 1929), 235-236.
43. *The Twenty-Fourth Year Book of the National Society for the Study of Education,* Part I. Bloomington, Illinois: Public School Publishing Company, 1925. 363 pp.
44. Vogel, Mabel; Jaycox, Emily; and Washburne, Carleton. "A Basic List of Phonics for Grades I and II," *Elementary School Journal,* XXII (February, 1932), 436-443.
45. Washburne, Carleton and Vogel, Mabel. "A Revised List of Phonics for Grade II," *Elementary School Journal,* XXVIII (June, 1928), 771-777.
46. Wheat, Harry G. *The Teaching of Reading.* Boston: Ginn and Company, 1923. ix, 346 pp.
47. Winch, William H. *Teaching Beginners to Read in England: Its Methods, Results, and Psychological Basis.* Journal of Educational Research Monographs, No. 8. Bloomington, Illinois: Public School Publishing Company, 1935. 186 pp.

DUKE UNIVERSITY RESEARCH STUDIES IN EDUCATION

Already published:

No. 1 The Effects of Unfamiliar Settings on Problem-Solving by W. A. Brownell and Lorena B. Stretch.

No. 2 Variability in Results from New-Type Achievement Tests by Earl V. Pullias.

No. 3 Learning as Reorganization: An Experimental Study in Third-Grade Arithmetic by W. A. Brownell with the assistance of Kenneth G. Kuehner and William C. Rein.

No. 4 Growth in Understanding of Geographic Terms in Grades IV to VII by T. J. Eskridge, Jr.

No. 5 The Effect of Varied Amounts of Phonetic Training on Primary Reading by Donald C. Agnew.

Date Due

378.756 D877R no.1-5 229853

Duke University Libraries
D01129207M

D01129207M
Duke Library Service Center